

Descriptive Statistics in Clinical Trials

Shu-Ping Chen

2020/10/30

Descriptive Statistics in Clinical Trials

BioMed DOE: RCBD

The effect of the drug d-amphetamine sulfate on the behavior of rats was the object of the experiment. The behavior under study was the rate at which water-deprived rats pressed a lever to obtain water. The response was the lever press rate defined as the number of lever presses divided by the elapsed time of the session. The treatment factor levels were five different dosages of the drug in milligrams per kilogram of body weight, including a control dosage consisting of saline solution. An experiment, or run, consisted of injecting a rat with a drug dosage, and after one hour an experimental session began where a rat would receive water each time after a second lever was pressed. An individual rat could be used in many experiments by repeatedly injecting it with different doses of the drug (after an appropriate washout period) and by observing the lever pressing behavior. Each rat received all five doses in a random order with an appropriate washout period in between. Data is shown on Table 1. Data is in file DOERCBDRatBehavior.csv,

Introduction

The purpose of this experiment is to explore the effect of d-amphetamine sulfate on the behavior of rats. The behavior studied is the rate at which a mouse presses a lever to obtain water when it lacks water, and the rate is defined as the number of lever presses divided by the elapsed time of the test. The drugs in this test are divided into five different doses of the same drug, and the fixed dose contained in the saline solution is also included in the five different doses of drugs.

The purpose of the study was to observe the effects of different doses of drugs on the behavior of rats. According to the Table in Appendix 1, the data includes the number of the tested mouse, the drug group, and the compression rate after the injection of the drug. There is no missing value in the data. The drug group is evenly distributed, all of which are 10 observations. Then observe after the injection of the drug, the distribution of the compression rate can be found to be roughly symmetrical, and it can also be known in the subsequent verification that the data conforms to the normal distribution.

Discussion

As can be seen from the following table, when the d-amphetamine sulfate has not been injected, the average compression rate of the tested mice is 0.764; when the injection dose is 0.5, the average value rises to 0.934; when the injection dose is 1, the average value continues to rise to 1.014; when the injection dose is 1.5, the average value drops slightly to 1.009; when the injection dose is 2, the average value drops to 0.85. According to the above statistics, the average compression rate of the rats after the injection of d-amphetamine sulfate was not lower than the average compression rate of the rats without injection of d-amphetamine sulfate.

Table 1: Mean and s.e of compression rate of different doses

dose	mean	sd
0	0.764	0.189
0.5	0.934	0.221
1	1.014	0.185
1.5	1.009	0.183
2	0.850	0.258

As can be seen from the figure below, among the five different doses of the same medicine, the average compression rate at a dose of 1 (mg/kg body weight) is the highest. When the dose is decreased or increased from 1 (mg/kg body weight), the average compression rate tends to decrease.

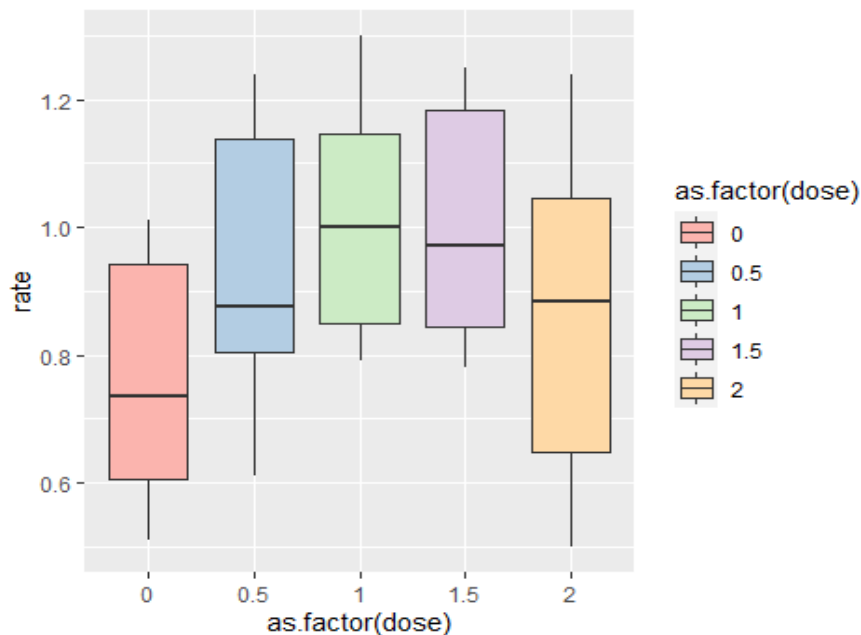


Figure 1: Compression rate of different doses Boxplot

From the figure below, we can clearly find the compression rate distribution of each observation when injecting different doses. Among them, the compression rate distribution at a dose of 2 (mg/kg body weight) is wider, while in the case of injecting other doses, The compression rate distribution of each observation value is more concentrated relative to the distribution of dose 2 (mg/kg body weight).

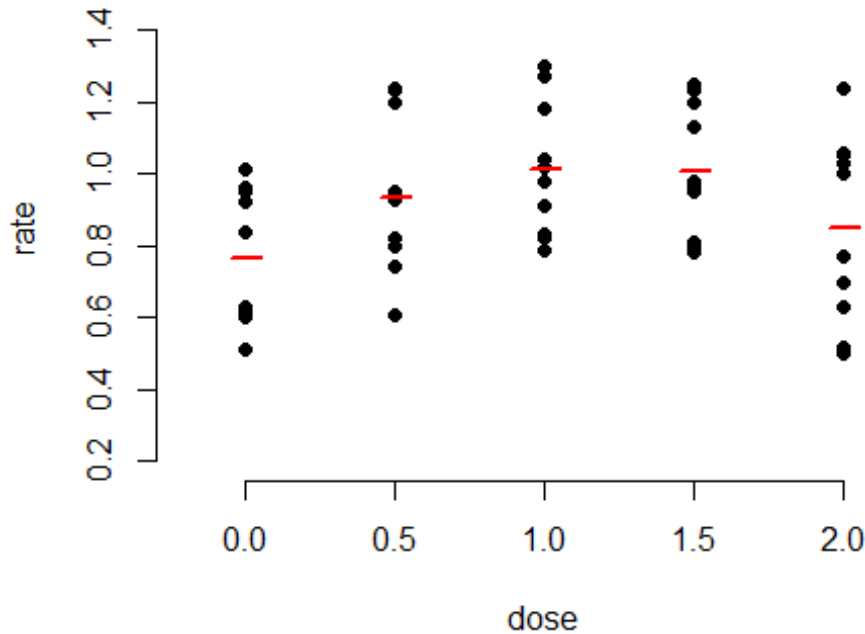


Figure 2: Compression rate of different doses Dot plot

Assumptions of one-way ANOVA is:

1. The sample is randomly sampled
2. Independent between groups
3. The variance between samples is homogeneity
4. Sample distribution is normal distribution

Therefore, we use the following test to test whether the sample meets the above assumptions to perform one-way ANOVA. First of all, we can know from the introduction of this experiment that the samples in this experiment are randomly distributed and independent of each other.

Next, we use the Shapiro-Wilk normality test to test for normality to observe whether the data is in line with normality, and to choose which homogeneity test to do. The test result shows that its $p\text{-value}=0.1>0.05$, so we do not reject the null

hypothesis. Therefore, it can be inferred that the data conform to the normal distribution.

```
##
## Shapiro-Wilk normality test
##
## data: drb$rate
## W = 1, p-value = 0.1
```

Finally, Bartlett's test is used to test whether the variance between samples is homogeneity, and from the previous test, it can be seen that the variables are all under the normal distribution. From the Bartlett test's p-value=0.8>0.05, it can be seen that the null hypothesis in the Bartlett's test is not rejected, so it can be inferred that the variance between samples meets the homogeneity hypothesis.

$$\begin{cases} H_0: \sigma_{0.5}^2 = \sigma_1^2 = \sigma_{1.5}^2 = \sigma_2^2 = \sigma_{2.5}^2 \\ H_1: \text{not } H_0 \end{cases}$$

```
##
## Bartlett test of homogeneity of variances
##
## data: rate by dose
## Bartlett's K-squared = 2, df = 4, p-value = 0.8
```

From the above test, it can be known that this data satisfies the hypothesis of one-way ANOVA. Otherwise, according to the one-way ANOVA table, in the case of significance level is equal to 0.05, different doses of d-amphetamine sulfate have a significant effect on the compression rate.

$$Y_{ij} = \mu + \tau_i + \epsilon_i, \quad i = 1, \dots, 5$$

Table 2: one-way ANOVA

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## dose           4    0.46   0.1151    2.63  0.047 *
## Residuals     45    1.97   0.0438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It can be seen from the regression model analysis table that the dose 1 (mg/kg body weight) and dose 1.5 (mg/kg body weight) are significantly different from other groups, and the estimated coefficient can be known that the compression of the dose 1 (mg/kg body weight) group is the largest, followed by the compression rate of the dose 1.5 (mg/kg body weight) group, and the smallest compression rate group is the dose 2 (mg/kg body weight) group. Therefore, it can be inferred that the compression rate increases as the dose increases until the dose 1.5 (mg/kg body weight), and when the dose reaches 2 (mg/kg body weight), the compression rate drops sharply.

$$y_i = \beta_0 + \beta_1 x_{ij} + \epsilon_{ij}, \quad i = 1, \dots, 50$$

x : dose; i : number of mice; j : number of doses group

Table 3: Summary of regression model analysis

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.764	0.066	11.550	0.000
dose0.5	0.170	0.094	1.817	0.076
dose1	0.250	0.094	2.672	0.010
dose1.5	0.245	0.094	2.619	0.012
dose2	0.086	0.094	0.919	0.363

Statistical model:

The following table is about the one-way ANOVA analysis, and the mean of estimated value of compression rate in this data is 0.9142, and the compression rate of mice, which don't get d-amphetamine sulfate is estimated to be -0.15; the compression rate effect of dose 0.5 is estimated to be 0.0198, dose; the compression rate effect of dose 1 is estimated to be 0.0998, the compression rate effect of dose 1.5 is estimated to be 0.0948, and the compression rate effect of dose 2 is estimated to be -0.0642.

Table 4: Summary of one-way ANOVA

dose	effect	UpperCLEffect	LowerCLEffect
0.0	-0.15	1.13	0.39
0.5	0.02	1.37	0.50
1.0	0.10	1.38	0.65
1.5	0.09	1.37	0.65
2.0	-0.06	1.36	0.34

Conclusion

When d-amphetamine sulfate has not been injected, the compression rate and the compression rate of dose 2 are faster, and if the dose is 0.5, 1, 1.5, the compression rate is slower.

BioMed DOE: Multiple LSQs

An experiment was designed to determine the effects of three diets on liver cholesterol in rats (A = control, B = control + vegetable fat, C = control + animal fat). Body weight classifications (H, M or L) of the rats and the litters from which they came were used to form a balanced set of Latin squares. The litter was nested in

squares (i.e., different litters were used in each square), whereas the weight classifications were not nested. The data is presented as follows. Data is in file DOEchomultiLSQ.csv.

Introduction

The main purpose of this test is to determine the effects of three diets on the liver cholesterol of mice (group A: control group, group B: control group + vegetable fat, C: control group + animal fat), and use the weight to classify the mice (H, M or L) and the nest they came from, form a set of balanced Latin squares. The rat's nest is nested in the grid, but the weight does not appear to be nested.

The purpose of this study is to observe whether the diet group of rats affects the value of liver cholesterol. According to the Table in Appendix II, this data contains 27 observations and five variables, which are the number of times the Latin square is repeated, the weight of the mouse, the litter from which the mouse comes, diet and liver cholesterol, and there are no missing values. The top four variables are categorical variables, and the number of each category is the same, and the distribution of liver cholesterol values is approximately normal.

Discussion

Through Table 5, It can be known that different diet groups have significant changes in liver cholesterol, so diet groups have a certain influence on liver cholesterol

Table 5: Variance analysis in two-stage nest design

Source	df	Type I SS	Mean square	F value	Pr > F
sq	2	0.040	0.020	3.83	0.05
litter(sq)	6	0.168	0.028	5.32	0.01
wt	2	0.196	0.098	18.65	<.001
treat	2	0.336	0.168	31.99	<.0001

It can be seen from the table below that the liver cholesterol of sq1 and sq2 is lower than the total average liver cholesterol; the liver cholesterol of sq3 is higher than the total average; the liver cholesterol of treat A and treat B is lower than the total average, while the liver cholesterol of treat C is lower than the total average. The liver cholesterol of the higher weight group is higher than the total average; the liver cholesterol of the middle weight and the lower weight group is lower than the total average; the liver cholesterol of litter1 in sq1, sq2, and sq3 is lower than the total average respectively, and litter3 also had higher liver cholesterol in sq1, sq2, and sq3. Among them, the liver cholesterol of litter2 in sq1 and sq2 was higher than the total average; on the contrary, the liver cholesterol in sq3 was lower than the total average.

Table 6: Summary of ANOVA model in nest design

variable	class	class2	effect	UpperCLEffect	LowerCLEffect
sq	1	.	-0.04	2.14	1.36
sq	2	.	-0.01	2.10	1.45
sq	3	.	0.05	2.17	1.51
treat	A	.	-0.08	2.05	1.36
treat	B	.	-0.08	1.91	1.52
treat	C	.	0.16	2.21	1.68
wt	H	.	0.11	2.23	1.57
wt	L	.	-0.09	2.00	1.39
wt	M	.	-0.02	2.08	1.46
litter	1	sq1	-0.13	2.01	1.24
litter	2	sq1	0.09	2.09	1.59
litter	3	sq1	0.03	2.29	1.27
litter	1	sq2	-0.10	1.77	1.59
litter	2	sq2	0.01	2.21	1.36
litter	3	sq2	0.08	2.24	1.49
litter	1	sq3	-0.04	2.30	1.30
litter	2	sq3	-0.05	2.07	1.51
litter	3	sq3	0.10	2.12	1.75

Conclusion



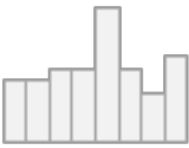
Through the above statistical analysis, it can be inferred that there are significant differences in the number of repetitions, the mouse's nest, the weight of the mouse, and the eating habits in the case where the nesting layer the mouse comes from is in the Latin square matrix. Among them, the liver cholesterol of the test mice that ate animal fat was higher, and the other two groups were lower; the hepatic cholesterol of the test mice in the higher body weight group was higher, and the other two groups were lower; one of the test mice from litter1 Liver cholesterol is lower, while the liver cholesterol of mice from litter3 is higher.

Appendix

Appendix Table 1: information of data in RCBD

Dimensions: 50 x 3



Duplicates: 0



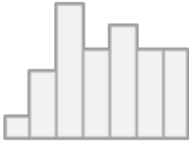
Variable	Stats / Values	Freqs (% of Valid)	Graph	Missing
rat [integer]	Mean (sd) : 5.5 (2.9) min < med < max: 1 < 5.5 < 10 IQR (CV) : 5 (0.5)	1 : 5 (10.0%) 2 : 5 (10.0%) 3 : 5 (10.0%) 4 : 5 (10.0%) 5 : 5 (10.0%) 6 : 5 (10.0%) 7 : 5 (10.0%) 8 : 5 (10.0%) 9 : 5 (10.0%) 10 : 5 (10.0%)		0 (0%)
dose [factor]	1. 0 2. 0.5 3. 1 4. 1.5 5. 2	10 (20.0%) 10 (20.0%) 10 (20.0%) 10 (20.0%) 10 (20.0%)		0 (0%)
rate [numeric]	Mean (sd) : 0.9 (0.2) min < med < max: 0.5 < 0.9 < 1.3 IQR (CV) : 0.3 (0.2)	38 distinct values		0 (0%)

Appendix Table 2: information of data in Multiple LSQs

Dimensions: 27 x 5

Duplicates: 0

Variable	Stats / Values	Freqs (% of Valid)	Graph	Missing
sq [integer]	Mean (sd) : 2 (0.8) min < med < max: 1 < 2 < 3 IQR (CV) : 2 (0.4)	1 : 9 (33.3%) 2 : 9 (33.3%) 3 : 9 (33.3%)		0 (0%)
wt [character]	1. H 2. L 3. M	9 (33.3%) 9 (33.3%) 9 (33.3%)		0 (0%)

litter [integer]	Mean (sd) : 2 (0.8)	1 : 9 (33.3%)		0 (0%)
	min < med < max:	2 : 9 (33.3%)		
	1 < 2 < 3	3 : 9 (33.3%)		
	IQR (CV) : 2 (0.4)			
treat [character]	1. A	9 (33.3%)		0 (0%)
	2. B	9 (33.3%)		
	3. C	9 (33.3%)		
cho [numeric]	Mean (sd) : 1.8 (0.2)	22 distinct values		0 (0%)
	min < med < max:			
	1.4 < 1.8 < 2.1			
	IQR (CV) : 0.3 (0.1)			